Early Evolution of Ionotropic GABA Receptors and Selective Regimes Acting on the Mammalian-Specific Theta and Epsilon Subunits

Christopher J. Martyniuk[1¤], Stéphane Aris-Brosou[1,2*], Guy Drouin[1], Joel Cahn[1], Vance L. Trudeau[1]

1 Department of Biology and Centre for Advanced Research in Environmental Genomics, University of Ottawa, Ottawa, Ontario, Canada, **2** Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada

Background. The amino acid neurotransmitter GABA is abundant in the central nervous system (CNS) of both invertebrates and vertebrates. Receptors of this neurotransmitter play a key role in important processes such as learning and memory. Yet, little is known about the mode and tempo of evolution of the receptors of this neurotransmitter. Here, we investigate the phylogenetic relationships of GABA receptor subunits across the chordates and detail their mode of evolution among mammals. *Principal Findings.* Our analyses support two major monophyletic clades: one clade containing GABA$_A$ receptor α, γ, and ε subunits, and another one containing GABA$_A$ receptor ρ, β, δ, θ, and π subunits. The presence of GABA receptor subunits from each of the major clades in the *Ciona intestinalis* genome suggests that these ancestral duplication events occurred before the divergence of urochordates. However, while gene divergence proceeded at similar rates on most receptor subunits, we show that the mammalian-specific subunits θ and ε experienced an episode of positive selection and of relaxed constraints, respectively, after the duplication event. Sites putatively under positive selection are placed on a three-dimensional model obtained by homology-modeling. *Conclusions.* Our results suggest an early divergence of the GABA receptor subunits, before the split from urochordates. We show that functional changes occurred in the lineages leading to the mammalian-specific subunit θ, and we identify the amino acid sites putatively responsible for the functional divergence. We discuss potential consequences for the evolution of mammals and of their CNS.

INTRODUCTION

Gene duplication followed by gene divergence is one of the major mechanisms responsible for the evolution of new functions [1]. It is thought to underlie the evolution of vertebrates and more particularly their complex and specialized central nervous system (CNS) [2]. Two classes of proteins that underwent such a mechanism are cationic (e.g., acetylcholine, serotonin) and anionic (e.g., glycine, γ-aminobutyric acid) ligand-gated channels, estimated to have diverged before the origin of eukaryotes [3]. However, little more is known about the mode and tempo of evolution of these receptors. Such knowledge may provide us with insights into the structural and functional complexity of these receptors in the central nervous system and in their role in mammalian evolution.

The γ-aminobutyric acid (GABA) is the major inhibitory neurotransmitter found in the vertebrate brain and is involved in CNS development and organization [4], neuroendocrine function [5], and neural processes such as learning and memory [6]. GABA is also present in the nervous system of non-vertebrate taxa, for example, flatworms [7], arthropods [8,9] and early chordates [10]. GABA synaptic transmission is achieved through membrane bound postsynaptic receptors. Currently, there are three major classes of GABA receptors identified in the mammalian CNS: GABA$_A$, GABA$_B$ and GABA$_C$. These are distinguished according to their composition, pharmacology and localization. Ionotropic GABA$_A$ receptors are ligand-gated chloride (Cl$^-$) channels consisting of both high abundance subunits (α1-6, β1-4, γ1-3, δ) and low abundance subunits (ε, θ, and π) [11]. Changes in the abundance and composition of these subunits have been shown to induce differences in GABA$_A$ receptor sensitivity and response [12,13]. The current structural model of the GABA$_A$ receptor is a pentameric receptor with binding sites for the GABA ligand and for receptor modulation by benzodiazepines, neurosteroids,

ethanol, and barbiturates [14]. Metabotropic GABA$_B$ receptors are members of the seven transmembrane domain family and are coupled to downstream calcium and potassium channels via G-proteins [15]. Finally, GABA$_C$ receptors are also pentameric ionotropic Cl$^-$ channels that show similar membrane topology as the GABA$_A$ receptors. However, GABA$_C$ receptors have unique functional and electrophysiological characteristics, including a slower Cl$^-$ conductance and insensitivity to bicuculline and other GABA$_A$ receptor modulators [14,16]. Ionotropic GABA$_C$ receptors are composed of ρ subunits that are highly expressed in the vertebrate retina and preferentially localized to bipolar cells [17] but are also found in the spinal cord and pituitary [18]. Both spatial and temporal regulation of GABA receptor subunit expression provide functional diversity to the GABA receptor family.

· ·

Fourteen of the human $GABA_A$ receptor genes are clustered on four chromosomes (chromosomes 4, 5, 15, and X: [19,20]). Two of these clusters contain two genes coding for receptor α subunits, one gene coding for a receptor β subunit and one gene coding for a receptor γ subunit whereas the other two clusters contain single genes coding for a receptor α, β and γ subunit–with the ε subunit gene replacing the γ subunit gene on the X chromosome [19]. Evidence based on chromosomal organization, intron/exon structure and transcriptional organization suggests that these four clusters originated from duplications of (and within) an ancestral $GABA_A$ receptor gene cluster containing single genes coding for a receptor α, β and γ subunit [19–21]. The other two $GABA_A$ receptor genes, coding for the receptor δ and the π subunits, also likely arose by duplication of an ancestral $GABA_A$ receptor gene(s) but are not part of the four $GABA_A$ receptor gene clusters [20]. $GABA_C$ receptor genes share only about 35% amino acid sequence identity with $GABA_A$ receptor genes [20]. This suggests that these two types of genes may have a common origin and have since acquired distinct functions [20,22]. However, these studies largely focused on mammalian sequences and the evolutionary history of ionotropic GABA receptor subunits before the origin of mammals is unclear.

Our goal here is twofold. First, we extend our knowledge of the phylogenetic relationships between ionotropic GABA receptor subunits, in particular by including recently available genomes such as those of the sea squirt *Ciona intestinalis* and the pufferfish *Takifugu rubripes*. Second, we test whether positive selection can explain the evolution of the mammalian-specific GABA receptor subunits θ and ε.

RESULTS

$GABA_A$ receptor phylogeny

Our estimate of the $GABA_A$ receptor phylogeny shows two major monophyletic clades (Fig. 1), which is consistent with previous studies [3,23–25]. This topology is robust to both the search algorithm used and the model of evolution. Indeed, the same topology was obtained by both maximum likelihood and Bayesian approaches, with the latter integrating over different substitution models (where the substitution model used in maximum likelihood [see Methods] had a posterior probability of one). The phylogenetic tree of vertebrate sequences containing the $GABA_A$ receptor α, γ, and ε subunits shows that these receptor subunits do not result from non-orthologous gene displacement [26] but appear to be derived from a common ancestor (Fig. 1 and 2). Subunits α are divided into two strongly supported groups: a clade composed of subunits α4 and α6 and the other composed of subunits α1-3 and α5. Subunits γ1 and γ2 form a strongly supported group, whereas the γ3, γ4 and ε subunits all group together. This suggests that the $GABA_A$ receptor ε subunits are derived from γ subunits. This is in agreement with previous suggestions that the $GABA_A$ receptor subunits γ4 of chicken and ε found in mammals are likely to be orthologous [20,27].

The relationship of the GABA receptor subunits involved in benzodiazepine binding (subunits ρ, β, δ, θ, and π) is uncertain and depends on whether invertebrate GABA/glycine-like receptor sequences are included or not in the analysis. Indeed, when the invertebrate GABA/glycine-like receptor sequences are included (81-sequence data set), subunits δ and π are found to be the sister group to the other GABA receptor sequences in this clade (Fig. 1). In contrast, the tree estimated without the invertebrate GABA/glycine-like receptor sequences (55-sequence data set) shows that subunits ρ are the sister group to other subunit families within this clade (Fig. 3). However and importantly, for both trees, the $GABA_C$ receptor ρ subunits, the $GABA_A$ receptor θ and β subunits, and the $GABA_A$ receptor π and δ subunits, form three strongly supported groups (Fig. 1 and 3).

Analysis of selective pressures

Our phylogenetic analyses suggested that the θ and ε paralogs have undergone a period of accelerated evolution following the duplication event. Such a period of accelerated evolution could be due to an episode of positive selection affecting these branches. To test this hypothesis, we used a codon substitution model, which measures selective pressures by estimating the nonsynonymous to synonymous rate ratio. This ratio is denoted ω, with ω = 1, <1 and >1 indicating neutral evolution, purifying and positive selection, respectively [28]. The comparison of a null codon model where selective pressures are constant both along lineages and among sites (H_0) against a branch-specific or "free-ratio" model (denoted Br in Tables 1 and 2) showed an extensive variation of the ω rate ratio among branches in both data sets (Tables 1 and 2).

Our results suggest that the $GABA_A$ receptor θ subunit was subjected to positive selection. For reference, we label branches and clades following the conventions set in Figure 4a. When allowing selective pressures to differ specifically in the branch leading to this θ subunit (branch number 5: model H_1), the model does not significantly explain the data better (at the 1% level) than the constant ω rate ratio model (H_0; Table 1). However, specifically allowing for a different rate in the branch of interest (branch number 5) and allowing ω to vary among sites within this branch suggested the existence of sites evolving under positive selection in this branch. Indeed, both the test of positive selection (model A; see Methods and [29]) and its variant (model B; see Methods and [30]) have highly significant *P*-values (Table 1). The estimate of ω for the sites allowed to be under positive selection in branch 5 or $\omega_5^{(2)}$ under model B (Table 1) is infinity because the synonymous rate along this branch is almost zero. Over all the branches, the average synonymous rate was .356 (excluding root branches), which suggests that saturation may not be an issue for these analyses. However, the sites estimated under models A and B are all different, except for site M274W. While both models consistently detect some evidence for positive selection, the identity of these sites is most likely unreliable, save for the possible exception of site M274W.

As of July 2007, no three-dimensional (3D) structures for GABA receptor subunits were deposited in the Protein DataBank (PDB: www.pdb.org). Similarity searches based on BLASTp using the protein sequence of the human θ paralog (accession number: AAI09211) returned no structure from PDB. Homology modeling with SWISS-MODEL [31] only produced a model between amino acid positions 291 and 347. This prediction was based on the NMR-obtained template whose PDB accession number is 1VRY, which corresponds to the second and third transmembrane domains of the α-1 subunit of human glycine receptor. The use of recently produced structure of the acetylcholine binding receptor (PDB: 1I9B; [32]) as a template failed to produce any 3D model: this is because the 17% similarity with the human θ paralog is below the 25% threshold of SWISS-MODEL. 3D-JIGSAW [33] produced a larger model that encompasses the previous one and spans from position 62 to 372 (Fig. 5). Because 3D-JIGSAW is more liberal than SWISS-MODEL, its prediction is expected to be somewhat inaccurate outside of the glycine receptor domain. To assess the robustness of our 3D model, we also produced 3D models based on the amino acid sequence of human paralogs β1 (CAA32875), β2 (AAB29370) and β3 (AAA52511). The models produced by 3D-JIGSAW had, after structural alignment, root mean square deviations of 1.21Å, 1.15Å and 1.60Å, respectively,
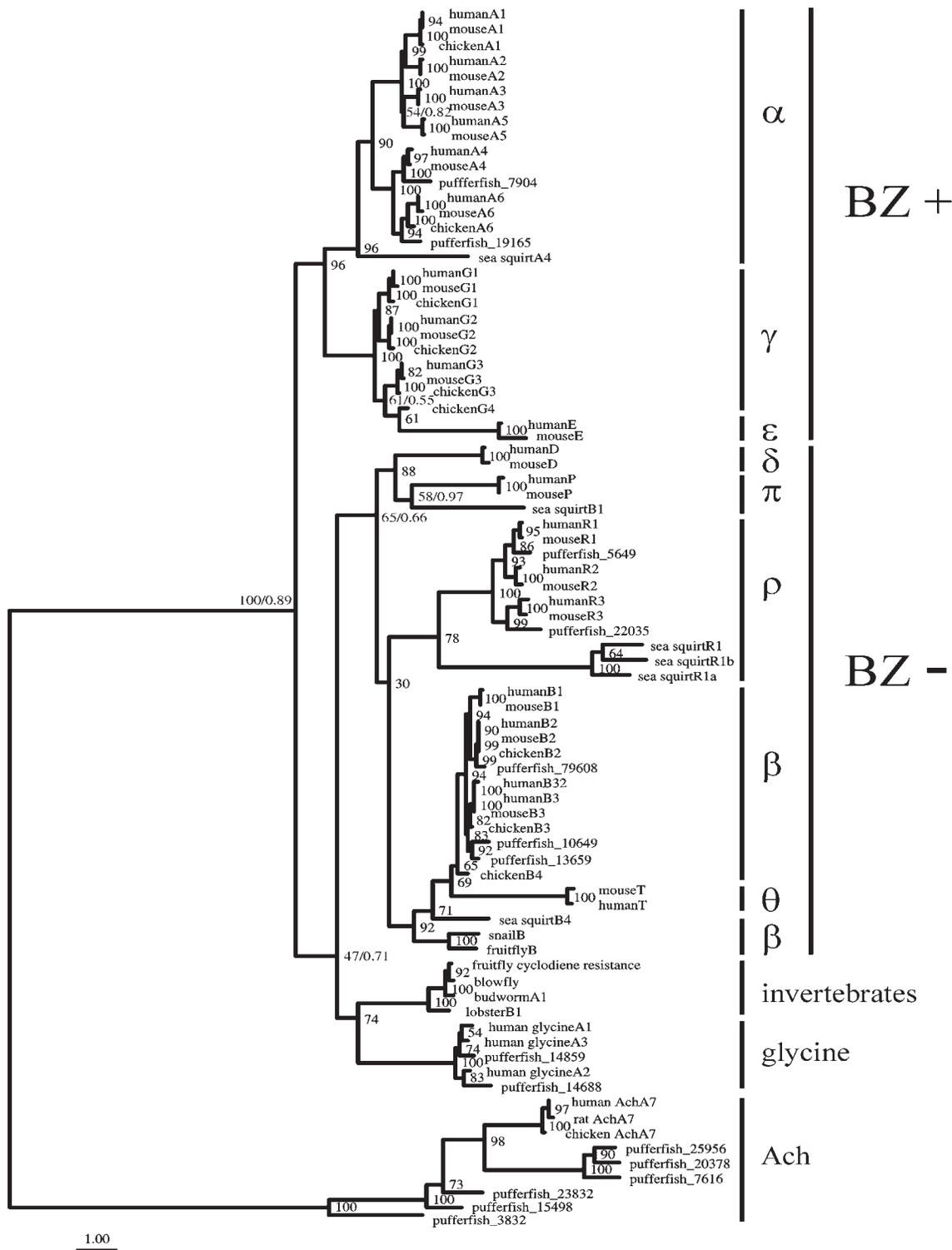
Figure 1. Phylogeny of GABA receptor protein sequences. The two clades, based on the presence (α, γ, and ε) or absence (ρ, β, δ, θ, and π) of a benzodiazepine binding site, are indicated by BZ+ and BZ–, respectively. Bootstrap values (maximum likelihood analysis) are indicated at all nodes while posterior probabilities (Bayesian analysis) are only indicated when smaller than 100%. The scale bar represents 1 amino acid substitutions per site. Ach: acetylcholine receptor.
doi:10.1371/journal.pone.0000894.g001

with the model produced with the human θ paralog. These figures are smaller than the 2.5Å resolution usually obtained by X-ray diffraction, so that our model is relatively robust to the protein sequence used within the β/θ group. Although the interpretation

of these models should be based on a more rigorous analysis, three domains are apparent on this model (Fig. 5): an N-terminal domain composed of essentially of β-sheets that might correspond to the ligand-binding domain, an intermediate domain composed

Figure 2. Phylogeny of vertebrate GABA receptor protein sequences with benzodiazepine binding sites. Bootstrap values are indicated at the nodes and the scale bar represents 1 amino acid substitutions per site.
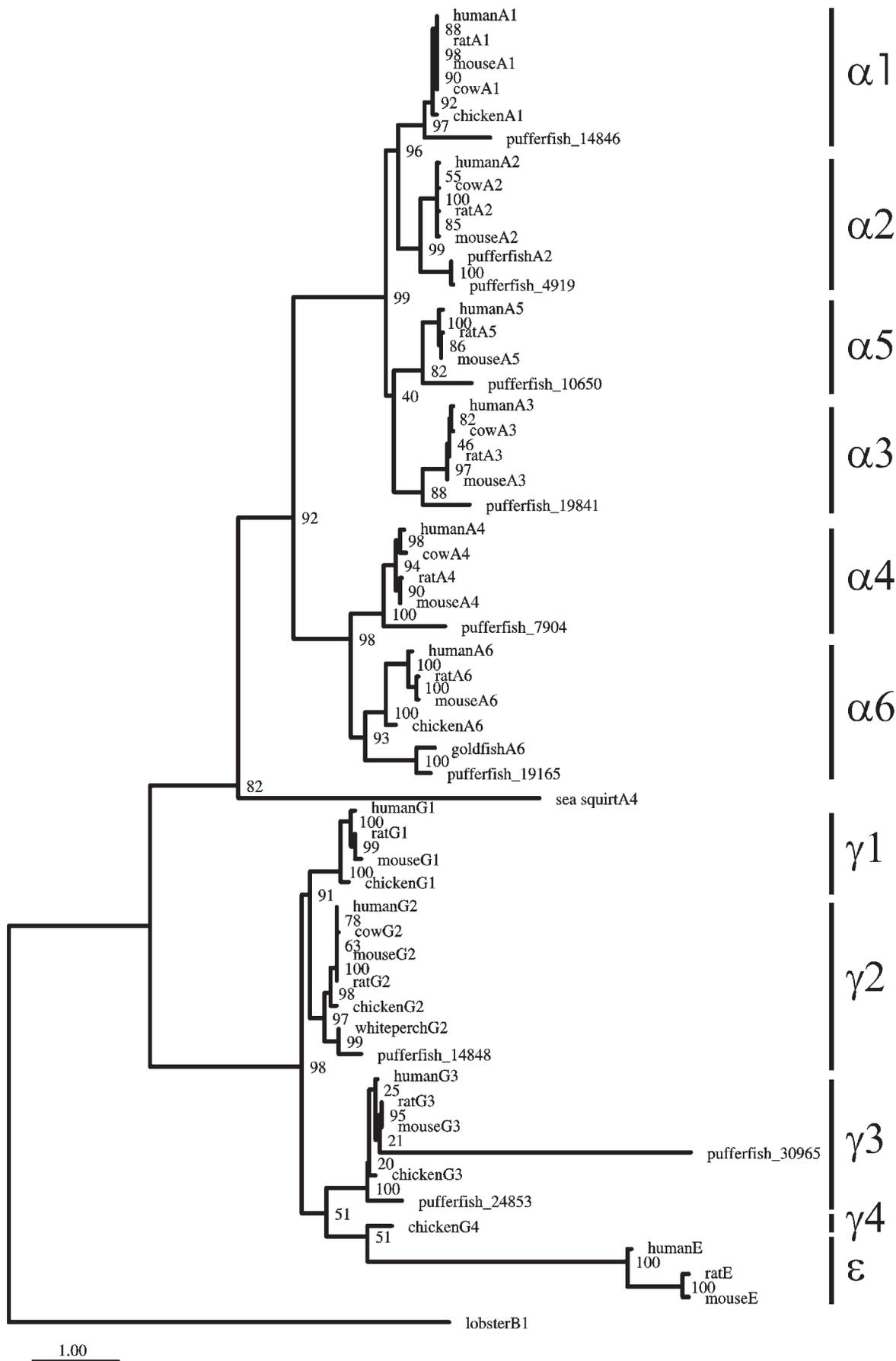doi:10.1371/journal.pone.0000894.g002

Figure 3. Phylogeny of vertebrate GABA receptor protein sequences without benzodiazepine binding sites. Bootstrap values are indicated at the nodes and the scale bar represents 1 amino acid substitutions per site.
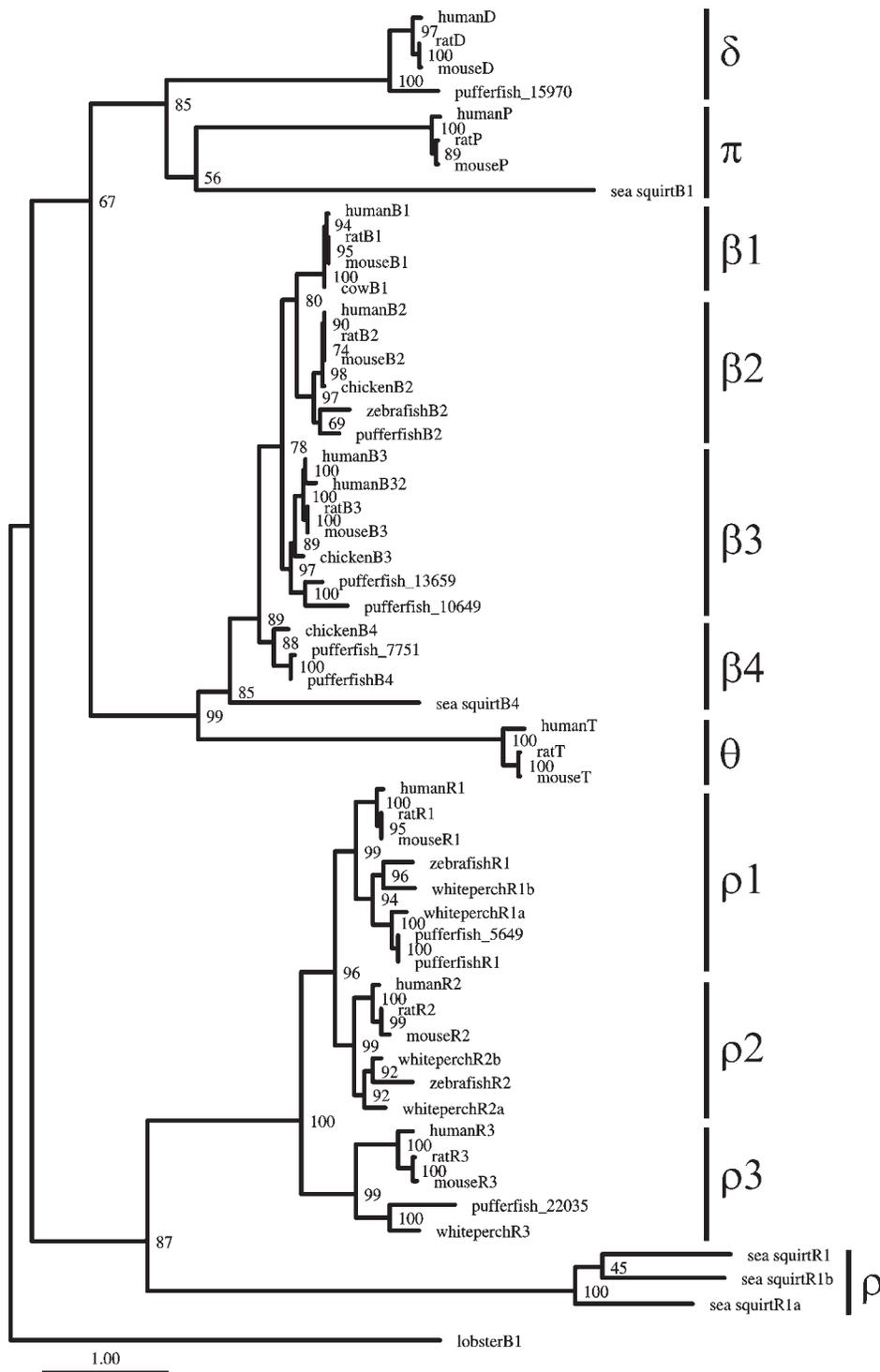doi:10.1371/journal.pone.0000894.g003

essentially of α-helices that might correspond to the transmembrane region and a C-terminal domain that might correspond to an intracellular region. This interpretation is consistent with what is known of our model's template (1VRY) and, more generally, of pentameric ligand-gated ion channels such as nAChR, 5-HT$_3$, GABA$_C$, and glycine receptors [34]. The sites putatively evolving under positive selection along branch 5 are located in all three

putative domains of the receptor (Fig. 5a–b). Note that site M274W, identified by both model A and model B is located within the putative transmembrane region on the first α-helices (M1: Fig.5; [34]). Further analyses allowing ω to differ among branches leading to the β paralogs (models H$_2$ and H$_3$) are consistent with an episode of positive selection leading to the θ paralogs, followed by a regime of negative selection. To summarize, evidence for

Table 1. Model comparisons and parameter estimates under models of constant (H$_0$) or variable ω rate ratios across branches (Br), clades (H$_1$ to H$_3$) or both clades and sites (A and B) for the θ subunit.

Model	ℓ	H$_A$	P-value	Parameter estimates	Positively selected sites
H$_0$	−8978.84	−	−	$\hat{\omega} = 0.0521$	None
Br	−8903.32	H$_0$	<0.0001	−	N/A
H$_1$	−8976.36	H$_0$	0.0262	$\hat{\omega}_0 = 0.0515$, $\hat{\omega}_5 = 0.4796$	N/A
A$_0$	−8958.66	−	−	$\hat{p}_0 = 0.46839$, $\hat{p}_1 = 0.02186$	N/A
				$\hat{\omega}_5^{(0)} = 0.04777$, $\omega_5^{(1)} = 1.00000$, $\omega_5^{(2)} = 1.00000$	
A	−8952.09	A$_0$	0.0003	$\hat{p}_0 = 0.44047$, $\hat{p}_1 = 0.02052$	(R24N), ({VA}135H), T143C, G224M, Y238R, P253I, M274W, K368P, (M458E), L571E, {DN}605P
				$\hat{\omega}_5^{(0)} = 0.04787$, $\hat{\omega}_5^{(1)} = 1.00000$, $\hat{\omega}_5^{(2)} = 2.22817$	
B$_0$	−8861.57	−	−	$\hat{p}_0 = 0.60788$, $\hat{\omega}_0 = 0.01456$, $\hat{\omega}_1 = 0.11773$	None
B	−8852.74	B$_0$	0.0001	$\hat{p}_0 = 0.26659$, $\hat{p}_1 = 0.14684$	M185L, M274W, M308L, Q363R
				$\hat{\omega}_5^{(0)} = 0.01043$, $\hat{\omega}_5^{(1)} = 0.13544$, $\hat{\omega}_5^{(2)} = \infty$	
H$_2$	−8924.96	H$_0$	<0.0001	$\hat{\omega}_0 = 0.0149$, $\hat{\omega}_4 = 0.1597$, $\hat{\omega}_5 = 0.6079$	N/A
H$_3$	−8922.42	H$_0$	<0.0001	$\hat{\omega}_0 = 0.0073$, $\hat{\omega}_1 = 0.0187$, $\hat{\omega}_2 = 0.0064$	N/A
				$\hat{\omega}_3 = 0.0204$, $\hat{\omega}_4 = 0.1598$, $\hat{\omega}_5 = 0.6783$	

Notes–ℓ: log-likelihood value; H$_A$: alternative hypothesis to the current model; ω$_0$: background rate; ω$_{1-5}$: branch/clade specific rates as indicated in Figure 4; p_{0-2}: proportions of sites in each rate category. The hat notation indicates parameters that are free to vary. Positively selected sites were identified with BEB and NEB between brackets for model A and NEB only for model B; a 99% cut-off level of posterior probability was used. Sites putatively under positive selection are numbered according to the human θ reference sequence (accession number BC109210/AAI09211). The underlined sites are those common to both model A and model B. Curly braces indicate equally parsimonious ancestral sites.

positive selection acting in this branch is strong, seems to have affected all three putative domains of the receptor subunit, but the identification of the actual sites should be taken with caution due to the relatively small number of sequences analyzed here.

Results from the ε data set only show evidence for relaxed selective pressures along the branch leading to the ε subunit. Here also, branches and clades are numbered following conventions in Figure 4b. As shown above, the reconstruction of this region of the GABA tree is unstable (Fig. 1 and 2). However, all results shown below were identical to the fifth decimal place under both tree reconstructions. This probably reflects (i) the low support for the ε clade (Fig. 1 and 2) and (ii) the relative robustness of site codon models to the exact topology of the tree as long as this tree is reasonable [35,36]. As with subunit θ, the fit of model H$_1$ is not significantly better than that of the one ratio model (H$_0$), but neither test with model A or with model B are significant (Table 2).

Table 2. Model comparisons and parameter estimates under models of constant (H$_0$) or variable ω rate ratios across branches (Br), clades (H$_1$ to H$_3$) or both clades and sites (A and B) for the ε subunit.

Model	ℓ	H$_A$	P-value	Parameter estimates	Positively selected sites
H$_0$	−9978.72	−	−	$\hat{\omega} = 0.0718$	None
Br	−9873.60	H$_0$	<0.0001	−	N/A
H$_1$	−9976.94	H$_0$	0.0592	$\hat{\omega}_0 = 0.0731$, $\hat{\omega}_5 = 0.0045$	N/A
M1a	−9932.48	−	−	$p_0 = 0.88903$, $\hat{\omega}_0 = 0.06590$	N/A
A$_0$	−9922.54	−	−	$\hat{p}_0 = 0.72680$, $\hat{p}_1 = 0.07894$	N/A
				$\hat{\omega}_5^{(0)} = 0.06560$, $\hat{\omega}_5^{(1)} = 1.00000$, $\hat{\omega}_5^{(2)} = 1.00000$	
A	−9921.96	M1a	<0.0001	$\hat{p}_0 = 0.74763$, $\hat{p}_1 = 0.08018$	{AG}206S, S231K
		A$_0$	0.2778	$\hat{\omega}_5^{(0)} = 0.06635$, $\hat{\omega}_5^{(1)} = 1.00000$, $\hat{\omega}_5^{(2)} = 3.18867$	
B$_0$	−9788.58	−	−	$\hat{p}_0 = 0.59972$, $\hat{\omega}_0 = 0.02012$, $\hat{\omega}_1 = 0.17322$	None
B	−9787.81	B$_0$	0.4615	$\hat{p}_0 = 0.58359$, $\hat{p}_1 = 0.38819$	(S231K, S330N, V351C) †
				$\hat{\omega}_5^{(0)} = 0.01985$, $\hat{\omega}_5^{(1)} = 0.17414$, $\hat{\omega}_5^{(2)} = 1.80401$	
H$_2$	−9909.20	H$_0$	<0.0001	$\hat{\omega}_0 = 0.0298$, $\hat{\omega}_4 = 0.2468$, $\hat{\omega}_5 = 0.0045$	None
H$_3$	−9896.02	H$_0$	<0.0001	$\hat{\omega}_0 = 0.0080$, $\hat{\omega}_1 = 0.0524$, $\hat{\omega}_2 = 0.0074$	None
				$\hat{\omega}_3 = 0.0306$, $\hat{\omega}_4 = 0.2475$, $\hat{\omega}_5 = 0.0045$	

Notes–ℓ: log-likelihood value; H$_A$: alternative hypothesis to the current model; ω$_0$: background rate; ω$_{1-5}$: branch/clade specific rates as indicated in Figure 4; p_{0-2}: proportions of sites in each rate category. The hat notation indicates parameters that are free to vary. Positively selected sites were identified with BEB and NEB between brackets for model A and NEB only for model B; a 95% cut-off level of posterior probability was used. Sites putatively under positive selection are numbered according to the human ε reference sequence (accession numbers HSU66661/AAB49284). †: these sites were putatively identified with a posterior probability >50% (see text). The underlined sites are those common to both model A and model B. Curly braces indicate equally parsimonious ancestral sites.
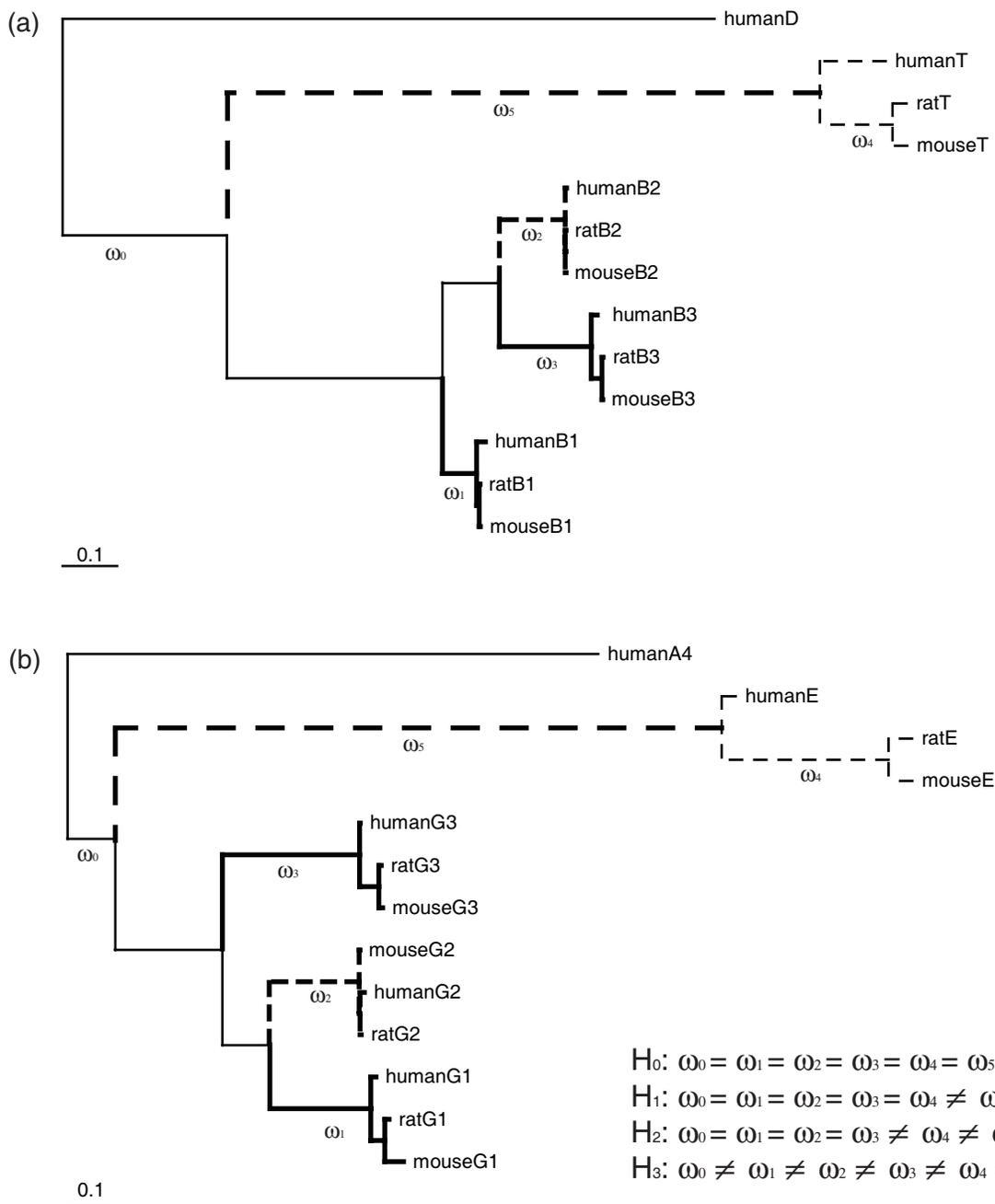
Figure 4. Phylogenetic trees used to formulate hypotheses about the selective forces that shaped the evolution of the GABA$_A$ subunits. (a) the GABA$_A$ receptor θ and β subunits ; (b) the GABA$_A$ receptor ε and γ subunits. Selective pressures are allowed to vary among sets of branches to test for the a priori hypotheses listed at the bottom of the figure: H$_0$: the null hypothesis that all branch-specific rates are equal; H$_1$: burst of evolution following the main gene duplication event (ε/γ and θ/β); H$_2$: also allows for a rate change after the burst of evolution; H$_3$: extends H$_2$ to allow for burst of evolution after all duplication events depicted on each tree. Hypotheses test are reported in Tables 1 and 2. The scale bar represents 0.1 substitutions per codon site, with branch lengths estimated under the null model (H$_0$).
doi:10.1371/journal.pone.0000894.g004

Importantly, the comparison of model A with a model called M1a [37] that assumes only two classes of sites, with $0<\omega<0$ and $\omega = 1$, is highly significant (Table 2). This latest comparison, called "test 1" in [29], cannot discriminate efficiently between positive selection and relaxed selective pressures [29]. Taken together, a significant "test 1" and a non-significant "test 2" suggest that the ε subunit only underwent relaxed constraints. Because of the conservative nature of "test 2" (see Methods and [29]), and because the average synonymous rate was 4.105 (excluding root branches) in the ε/γ

data set, more sequences will be required to definitely rule out the action of positive selection during the evolution of the ε subunit.

DISCUSSION
The present analyses and previous work [3,23] all suggest that an ancestral GABA-like receptor subunit gave rise to two mono-phyletic clades, categorized as subunits involved (α, γ, and ε) or not involved (ρ, β, δ, θ, and π) in benzodiazepine binding. As previously found [24,25,38], the presence of putative GABA α-, β-
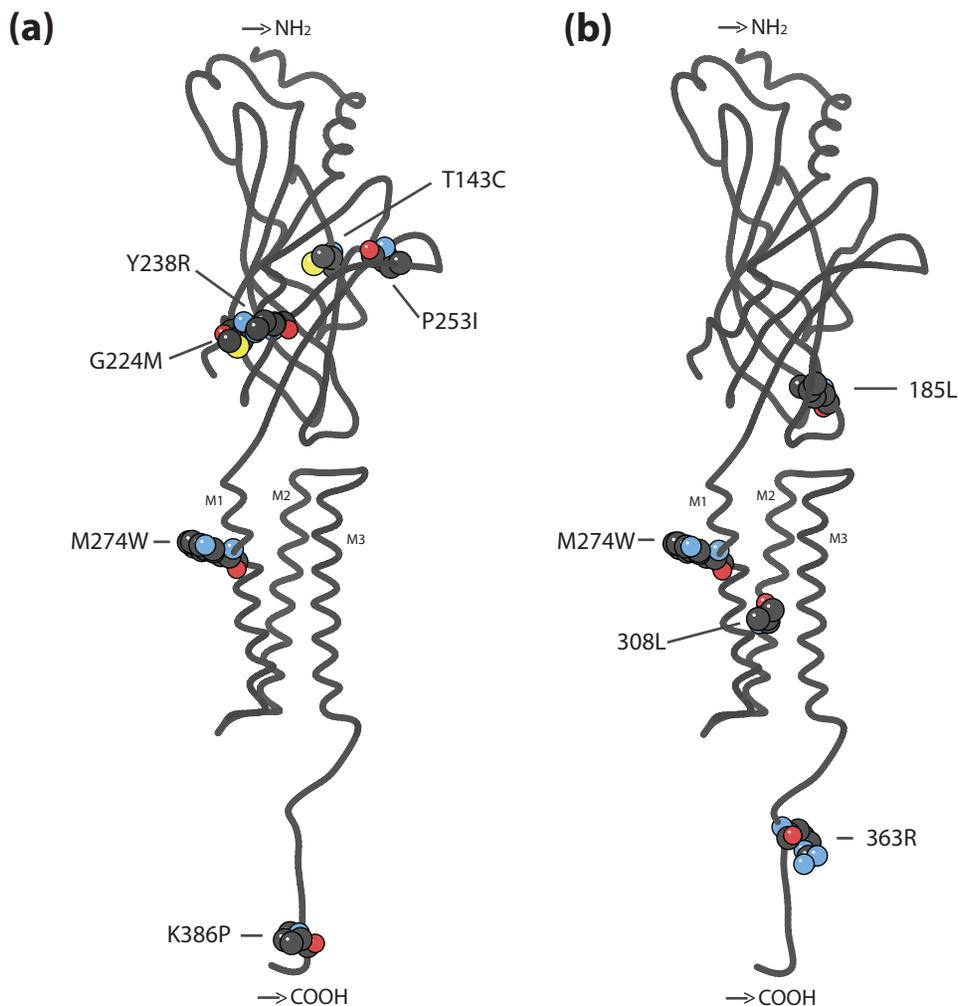
Figure 5. Mapping of the sites putatively under positive selection on the three-dimensional models of the θ subunit of the GABA_A receptor.
Sites putatively detected to be under positive selection on subunit θ under: (a) model A and (b) model B.
doi:10.1371/journal.pone.0000894.g005

and ρ-like receptor subunits in the genome of *C. intestinalis* support that this ancestral duplication most likely occurred before the divergence of urochordates. These results suggest that benzodiazepine sensitivity evolved early, in marked contrast to previous proposals [39,40]. In line with our finding, recent electrophysiological data suggest that invertebrates (e.g., *Hydra vulgaris*) are responsive to benzodiazepine modulation and this response is similar to the response to GABA [41]. On the other hand, the presence of α and β-like subunits in *C. intestinalis* does not necessarily indicate benzodiazepine sensitivity.

It was previously suggested that the GABA_A receptor δ subunit is the most primitive subunit within the GABA_A clade not involved in benzodiazepine binding [3]. Our phylogenetic analyses were unable to confidently determine the probable progenitor for receptor subunits in this clade (Fig. 1 and 3). Within the GABA_A receptor β subunit clade, it is interesting to note that our genome searches did not identify any β1 sequences in pufferfish or chicken genomes. This suggests that the pufferfish and chicken lineages may have independently lost their β1 paralog.

Our phylogenetic analyses also indicate that the mammalian-specific GABA_A subunits ε are derived from the γ4 subunit (Fig. 1 and 2), which is consistent with their chromosomal organization: Human GABA_A receptor ε genes are positioned on a location of the

X chromosome that corresponds to the position of GABA_A receptor γ4 genes on human chromosomes 4, 5 and 15 [19,20,27]. The higher rate of evolution of the GABA_A receptor ε subunits may explain why these subunits have so far only been found in mammals while the receptor γ4 subunits are present in both birds and reptiles [20]. Similarly, chromosomal location [20], intron-exon organization [21] and our phylogenetic analyses are consistent with the suggestion that the mammalian-specific GABA_A receptor θ subunits are derived from GABA_A receptor β4 subunits (Fig. 1 and 3).

Although we could only identify a small number of GABA_A receptor sequences from the *C. intestinalis* genome, the pufferfish genome contains orthologs to most of the GABA_A receptor sequences found in mammalian genomes. Furthermore, all the vertebrate genomes queried, including that of pufferfish, contain a single gene coding for most of the 20 GABA_A receptor subunits (α1-6, γ1-3, ε, δ, π, ρ1-3, β1-4 and θ). However in pufferfish, five of these gene families contain two paralogous gene copies instead of a single copy. Three of these pairs (pufferfish α2, β4 and ρ1) are composed of very similar genes that have undergone recent duplication events. It would be interesting to test experimentally whether the subunits encoded by these gene copies have similar functions. Examples of recent gene duplication events show that it is not necessarily the case: receptors for proglucagon-derived

peptides for instance exhibit differences in ligand-binding capabilities between recent paralogs [42]. The presence of two paralogous gene copies in some of the GABA_A receptor subunit families found in the pufferfish genome is consistent with a complete genome duplication event that occurred in the pufferfish lineage [43,44]. Further studies should elucidate whether duplicated GABA receptor subunits have a functional significance for GABAergic transmission in fish.

Our codon analyses showed that both GABA_A receptor subunits θ and ε experienced positive selection and relaxed constraints, respectively (Table 1–2). We further identified the putative amino acid sites that may be responsible for the functional divergence of subunit θ (Table 1 and Fig. 5). One position in subunit θ (274) was consistently identified as being under positive selection by different models. This position is located in putative transmenbrane segment M1, which, like M3, is though to interact with their neighboring subunits in the pentameric receptor [34]. The functional significance both of this location and of the adaptive substitution by a larger and aromatic residue at this position in subunit θ are unclear. Yet, positive selection acting on the sites in the N-terminal domain might have affected the ligand-binding affinity of GABA agonists and or antagonists, and maybe also allosteric transitions between different conformational states, that are the functions associated to this domain in nicotinic receptors [34]. On the other hand, positive selection acting on the C-terminal domain might have affected the sensitivity of the receptor, a property that is determined by the five loops constituting the pentameric receptor, at least in nicotinic receptors [34].

Relaxed constraints in the ε subunit would be consistent with the observation that rodent ε subunit have acquired an unusually large insertion of 483 amino acids in their second exon [27]. These subunits are expressed in the CNS, and are less abundant than other subunits [11]. In rats, the expression of ε subunits is associated with peptidergic neurons, such as those producing orexin, oxytocin, and gonadotropin-releasing hormone (GnRH). This suggests that GABA_A receptors with ε subunits might have a role in neuroendocrine function, such as that involved in the control of feeding and reproduction [45]. A recent study showed that ε subunits could increase GABA sensitivity up to 100-fold in *Xenopus* oocytes [46]. It was also shown that receptors containing the ε subunit could be insensitive to the GABA receptor modulators pregnanolone and pentobarbital [46]. Interestingly, these GABA receptor subtypes are predominant in the locus coeruleus, a nucleus in the brain stem that contains a large population of noradrenergic positive neurons. When this region is lesioned in rats, there is a disruption in the preovulatory surge of luteinizing hormone (LH), follicle-stimulating hormone (FSH), and prolactin [47] and a significant reduction in circulating plasma LH [48]. Further studies should assess the role of the GABA_A receptor θ and ε subunits in this neuroendocrine pathway. Should these subunits be involved in this pathway, our results would suggest that the rapid divergence of the GABA_A receptor θ and ε subunits played a role in the evolution of the neuroendocrine system in mammals. Recent studies utilizing point-mutations have shown that single amino acid changes in a GABA receptor subunit will have dramatic effects on the kinetics of the receptor [49,50]. Future studies are needed to determine whether or not these amino acid changes confer significant alterations in GABA receptor kinetics and function.

Conclusions
To conclude, our results show that (1) the two major clades of ionotropic GABA receptors arose before the split from urochordates, (2) the GABA_A receptor family evolved by both gains and losses of subtypes (e.g., teleost $\beta 4$, chicken $\gamma 4$, mammalian ε and θ) and (3) the

function of the GABA receptor subunits might have changed adaptively in the mammalian-specific GABA_A subunit θ, while relaxed constraints acted on subunit ε. These changes of selective regime might have played a role in the evolution of neuroendocrine functions controlling feeding and reproduction in mammals. We caution however that further research should be performed to experimentally test these functional divergence hypotheses.

MATERIALS AND METHODS
Genome database searches
Homologous gene queries were performed on the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov) server. BLASTn and BLASTp searches [51] were used to find homologous genes of full-length GABA receptor subunits α1-6, β1-3, γ1-3, δ, ε, θ, and π present in the completed rat genome project. The *Ciona intestinalis* sequences were obtained from the Joint Genome Institute (http://genome.jgipsf.org/ciona4/ciona4.home.html). Sequence alignments are available as supplementary Text S1 and S2.

GABA receptor phylogeny
We obtained and aligned a data set of 81 homologous protein sequences of the GABA receptor subunits queried above. This data set contains almost exclusively complete protein sequences with 997 aligned amino acid positions; this alignment is much longer than the actual length of GABA receptors (ca. 470 amino acids) because of the presence of large indels in some specific subunits. All aligned positions were conserved to help tease apart highly conserved proteins. To test for robustness of these analyses, we increased the species sampling of receptors involved in benzodiazepine binding (BZ+: α, γ, ε) and those that are not (BZ–: ρ, β, δ, θ, and π) [3,23]. Note that this BZ+/BZ– classification of subunits is somewhat artificial as two subunits are actually required to form a benzodiazepine sites [11]. We sampled 55 chordates including sea squirt and vertebrates for the BZ+ GABA_A receptor protein sequences (1,007 aligned amino acid positions) and 55 chordates including sea squirt and vertebrate BZ– protein sequences (795 aligned amino acid positions). Again, discrepancies between actual protein lengths and length of the alignment reflect the presence of indels.

Phylogenetic trees were inferred using both the maximum likelihood approach as implemented in PHYML 2.4.4 [52] and the Bayesian approach implemented in MrBayes 3.1.1 [53]. In both approaches, the 81-sequence data set was rooted with acetylcholine receptor protein sequences [3,23]. Based on the results from the analysis of this data set (Fig. 1), the BZ+ and BZ– data sets (Fig. 2 and 3, respectively) were rooted with the lobster $\beta 1$ (accession number AY098945) because this sequence provides a closer outgroup to both data sets (and that closer outgroups minimize possible errors in phylogenetic trees). ProtTest [54] was used to determine, based on the Akaike Information Criterion (minAIC), that the JTT+I+γ model of amino acid substitution [55,56] was the most appropriate model of evolution for this data set ($\hat{p}_I = 0.01$; $\hat{\alpha} = 0.819$); the second to most appropriate model was JTT+γ ($\Delta AIC = 48.81$), while the second to most appropriate rate matrix was WAG (+I+γ: $\Delta AIC = 444.74$; +γ: $\Delta AIC = 523.11$). Among-site rate variation modeled by a discrete γ distribution with eight rate categories. This maximum likelihood model was fitted independently to each data set.

Because model choice as performed by ProtTest can lead to underestimating uncertainty, the maximum likelihood analyses were complemented by a Bayesian analysis that integrates over different models of evolution: rather than selecting an a priori empirical model of substitution, a reversible-jump Markov chain Monte Carlo (RJ-MCMC) was constructed to integrate over

model uncertainty [57]. Models with equal prior probability were: Poisson, JTT, Dayhoff, MtREV, MtMAMM, WAG, rtREV, CpREV, VT and BLOSUM62 as described and implemented in MrBayes [53]. Among-site rate variation was modeled using a discrete γ distribution with five rate categories [56] plus a class of invariable sites. Under this mixed model of protein evolution, four independent RJ-MCMC samplers were run for ten million steps using different starting values. To decrease autocorrelation of the samples taken from the target distribution, steps along the chain were sampled every 1,000 accepted steps, a method known as thinning (e.g., [58]). To improve mixing, each sampler consisted of four chains, three of which were heated to different temperatures (e.g., [56]). By raising the likelihood function to a power <1, deep valleys of the likelihood surface become shallow, which facilitates their crossing by the sampler and hence improves the ability of the chain to explore the entire parameter space efficiently, in proportion of the target density. Sampling was realized from the non-heated chain. Burn-in length and convergence of the samplers were checked by plotting time series plots and checking that average standard deviations of split frequencies were lower than 0.01 [53]. The chains appeared to have converged by 10,000 steps; to be conservative, 100,000 steps were discarded as a burn-in.

Analysis of selective pressures

Codon data were obtained and split into two smaller data sets, one for each of the paralogous clade of interest: θ and ε. Each data set contained 13 sequences human (*Homo sapiens*), rat (*Rattus norvegicus*) and mouse (*Mus musculus*) copies of θ and ε paralogs, as well as the corresponding members of the group in which this clade was located: β for θ; γ for ε. Each tree was rooted by the closest human paralog (δ and α, respectively). Sites with ambiguous data were removed. A statistical approach was then used to detect functional divergence at individual codon sites within the pre-specified branches based on a procedure similar to that by Bielawski and Yang [59]: for each paralog the null hypothesis H_0 was that of no variable selective pressure among branches and among sites (Fig. 4). This hypothesis was contrasted by means of likelihood ratio tests against three potential alternative modes of evolution. In H_1, only the branch leading to the θ or to the ε clade was allowed to evolve at a different ω rate ratio. This branch (branch number five) is said to have a foreground rate, while all the other branches have the same background rate. In a second model, H_2, the θ or the ε clade was allowed to evolve at a rate that differs from both the foreground and the background rates. Finally in H_3, each paralog was allowed to evolve at its own rate after duplication.

In some cases, only a few sites are affected by an episode of positive selection within a given branch. A more powerful approach to detecting sites undergoing positive selection in such cases is to allow ω to vary among sites within the branch of interest using the "test of positive selection" or "test 2" described in [29] and implemented in PAML version 3.15 [60]. These "branch-site" models allow the ω ratio to vary both among branches and among the sites in the foreground branch. Model A has four classes of sites: class 0 includes conserved codons with $0 < \omega_5^{(0)} < 1$ is estimated from the data; class 1 includes codons that evolve neutrally ($\omega_5^{(1)} = 1$); classes 2a and 2b include codons that are conserved or neutral on the background branches, but are under

positive selection on the foreground branch, with $\omega_5^{(2)} > 1$ estimated from the data. The "test of positive selection" compares this model against a simpler (null) model, that does not allow for positive selection ($\omega_5^{(2)} = 1$), so that this model estimates one fewer parameters than its alternative. To be conservative, we used χ_1^2 as an approximation to the distribution of the test statistic under the null [29]. Robustness of these branch-site codon models was assessed by using a second test based on "model B" as described in [29], which is identical to model A described above except that $\omega_5^{(1)}$ is free to vary. Model B is compared to model M3 [35] which is a site model with two discrete rate categories. Sites putatively under adaptive evolution in the "test of positive selection" were identified by the Bayes empirical Bayes (BEB) procedure [37], that improves on the naïve empirical Bayes approach (NEB: [61]) by accommodating uncertainties of the maximum likelihood estimates; with "model B", only BEB is implemented to assign individual sites to rate categories. All codon models were run at least twice to check convergence. Ancestral amino acid residues were determined by parsimony (for Tables 1 and 2).

Three-dimensional (3D) structure predictions were carried out with 3D-JIGSAW [33] and with SWISS-MODEL [31]. Both tools predict structures by homology modeling, a technique that can be decomposed into five steps: (i) a query or parts thereof are aligned to one or more template protein sequences, as determined by BLASTp searches; templates must have a resolved 3D structure; (ii) these 3D segments are put together to form a preliminary model; (iii) side chains are adjusted to account for substitutions between the query and the templates; (iv) the model is examined for potential collisions between atoms; finally (v) the model is refined by limited energy minimization. The major inaccuracies of homology modeling usually stem from two sources: low sequence similarity and improper template selection [62]. These two sources of inaccuracies were assessed as explained in the text. Structural alignments of the backbones of the models and computation of root mean square deviations or RMSDs were performed with DeepView, which is available through the ExPASy Web site.

Author Contributions

Conceived and designed the experiments: SA GD VT. Performed the experiments: CM JC. Analyzed the data: SA CM GD. Wrote the paper: SA CM GD JC VT.

REFERENCES

1. Ohta T (1989) Role of gene duplication in evolution. Genome 31: 304–310.
2. Shimeld SM, Holland PW (2000) Vertebrate innovations. Proc Natl Acad Sci USA 97: 4449–4452.
3. Ortells MO, Lunt GG (1995) Evolutionary history of the ligand-gated ion-channel superfamily of receptors. Trends Neurosci 18: 121–127.
4. Kellogg CK (1998) Early developmental modulation of GABA$_A$ receptor function. Influence on adaptive responses. Perspect Dev Neurobiol 5: 219–234.

5. Arvat E, Giordano R, Grottoli S, Ghigo E (2002) Benzodiazepines and anterior pituitary function. J Endocrinol Invest 25: 735–747.

6. Myhrer T (2003) Neurotransmitter systems involved in learning and memory in the rat: a meta-analysis based on studies of four behavioral tasks. Brain Res Brain Res Rev 41: 268–287.

7. Eriksson KS, Panula P (1994) gamma-Aminobutyric acid in the nervous system of a planarian. J Comp Neurol 345: 528–536.

8. Orona E, Battelle BA, Ache BW (1990) Immunohistochemical and biochemical evidence for the putative inhibitory neurotransmitters histamine and GABA in lobster olfactory lobes. J Comp Neurol 294: 633–646.

9. Darlison MG (1992) Invertebrate GABA and glutamate receptors: molecular biology reveals predictable structures but some unusual pharmacologies. Trends Neurosci 15: 469–474.

10. Anadon R, Adrio F, Rodriguez-Moldes I (1998) Distribution of GABA immunoreactivity in the central and peripheral nervous system of amphioxus (*Branchiostoma lanceolatum* Pallas). J Comp Neurol 401: 293–307.

11. Whiting PJ (1999) The GABA-A receptor gene family: new targets for therapeutic intervention. Neurochem Int 34: 387–390.

12. Rabow LE, Russek SJ, Farb DH (1995) From ion currents to genomic analysis: recent advances in GABA$_A$ receptor research. Synapse 21: 189–274.

13. Costa E, Auta J, Grayson DR, Matsumoto K, Pappas GD, et al. (2002) GABA$_A$ receptors and benzodiazepines: a role for dendritic resident subunit mRNAs. Neuropharmacology 43: 925–937.

14. Bormann J (2000) The 'ABC' of GABA receptors. Trends Pharmacol Sci 21: 16–19.

15. Couve A, Moss SJ, Pangalos MN (2000) GABA$_B$ receptors: a new paradigm in G protein signaling. Mol Cell Neurosci 16: 296–312.

16. Bormann J, Feigenspan A (1995) GABA$_C$ receptors. Trends Neurosci 18: 515–519.

17. Koulen P, Brandstatter JH, Kroger S, Enz R, Bormann J, et al. (1997) Immunocytochemical localization of the GABA(C) receptor rho subunits in the cat, goldfish, and chicken retina. J Comp Neurol 380: 520–532.

18. Johnston GA (2002) Medicinal chemistry and molecular pharmacology of GABA(C) receptors. Curr Top Med Chem 2: 903–913.

19. Russek SJ (1999) Evolution of GABA(A) receptor diversity in the human genome. Gene 227: 213–222.

20. Darlison MG, Pahal I, Thode C (2005) Consequences of the evolution of the GABA(A) receptor gene family. Cell Mol Neurobiol 25: 607–624.

21. Simon J, Wakimoto H, Fujita N, Lalande M, Barnard EA (2004) Analysis of the set of GABA(A) receptor genes in the human genome. J Biol Chem 279: 41422–41435.

22. Cutting GR, Curristin S, Zoghbi H, O'Hara B, Seldin MF, et al. (1992) Identification of a putative gamma-aminobutyric acid (GABA) receptor subunit rho2 cDNA and colocalization of the genes encoding rho2 (GABRR2) and rho1 (GABRR1) to human chromosome 6q14-q21 and mouse chromosome 4. Genomics 12: 801–806.

23. Xue H (1998) Identification of major phylogenetic branches of inhibitory ligand-gated channel receptors. J Mol Evol 47: 323–333.

24. Dent JA (2006) Evidence for a diverse Cys-loop ligand-gated ion channel superfamily in early bilateria. J Mol Evol 62: 523–535.

25. Tsang SY, Ng SK, Xu Z, Xue H (2007) The evolution of GABA$_A$ receptor-like genes. Mol Biol Evol 24: 599–610.

26. Galperin MY, Koonin EV (1999) Functional genomics and enzyme evolution. Homologous and analogous enzymes encoded in microbial genomes. Genetica 106: 159–170.

27. Sinkkonen ST, Hanna MC, Kirkness EF, Korpi ER (2000) GABA(A) receptor epsilon and theta subunits display unusual structural variation between species and are enriched in the rat locus ceruleus. J Neurosci 20: 3588–3595.

28. Goldman N, Yang Z (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol 11: 725–736.

29. Zhang J, Nielsen R, Yang Z (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol 22: 2472–2479.

30. Yang Z, Nielsen R (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol 19: 908–917.

31. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22: 195–201.

32. Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van Der Oost J, et al. (2001) Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. Nature 411: 269–276.

33. Bates PA, Kelley LA, MacCallum RM, Sternberg MJ (2001) Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins Suppl 5: 39–46.

34. Corringer PJ, Le Novere N, Changeux JP (2000) Nicotinic receptors at the amino acid level. Annu Rev Pharmacol Toxicol 40: 431–458.

35. Yang Z, Nielsen R, Goldman N, Pedersen AM (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155: 431–449.

36. Kosakovsky Pond SL, Frost SD (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 22: 1208–1222.

37. Yang Z, Wong WS, Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol 22: 1107–1118.

38. Okamura Y, Nishino A, Murata Y, Nakajo K, Iwasaki H, et al. (2005) Comprehensive analysis of the ascidian genome reveals novel insights into the molecular evolution of ion channel genes. Physiol Genomics 22: 269–282.

39. Nielsen M, Braestrup C, Squires RF (1978) Evidence for a late evolutionary appearance of brain-specific benzodiazepine receptors: an investigation of 18 vertebrate and 5 invertebrate species. Brain Res 141: 342–346.

40. Hebebrand J, Friedl W, Breidenbach B, Propping P (1987) Phylogenetic comparison of the photoaffinity-labeled benzodiazepine receptor subunits. J Neurochem 48: 1103–1108.

41. Kass-Simon G, Pannaccione A, Pierobon P (2003) GABA and glutamate receptors are involved in modulating pacemaker activity in hydra. Comp Biochem Physiol A Mol Integr Physiol 136: 329–342.

42. Irwin DM, Wong K (2005) Evolution of new hormone function: loss and gain of a receptor. J Hered 96: 205–211.

43. Hoegg S, Brinkmann H, Taylor JS, Meyer A (2004) Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J Mol Evol 59: 190–203.

44. Vandepoele K, De Vos W, Taylor JS, Meyer A, Van de Peer Y (2004) Major events in the genome evolution of vertebrates: paranome age and size differ considerably between ray-finned fishes and land vertebrates. Proc Natl Acad Sci USA 101: 1638–1643.

45. Moragues N, Ciofi P, Lafon P, Tramu G, Garret M (2003) GABA$_A$ receptor epsilon subunit expression in identified peptidergic neurons of the rat hypothalamus. Brain Res 967: 285–289.

46. Ranna M, Sinkkonen ST, Moykkynen T, Uusi-Oukari M, Korpi ER (2006) Impact of epsilon and theta subunits on pharmacological properties of alpha3beta1 GABA$_A$ receptors expressed in Xenopus oocytes. BMC Pharmacol 6: 1.

47. Anselmo-Franci JA, Franci CR, Krulich L, Antunes-Rodrigues J, McCann SM (1997) Locus coeruleus lesions decrease norepinephrine input into the medial preoptic area and medial basal hypothalamus and block the LH, FSH and prolactin preovulatory surge. Brain Res 767: 289–296.

48. Rocha AA, Guerra-Sa R, Silveira NA, Anselmo-Franci JA, Franci CR (2006) Neuropeptide Y in the medial basal hypothalamus and medial preoptic area during the induction of LH surge may be controlled by locus coeruleus. Neuropeptides 40: 57–63.

49. Derry JM, Dunn SM, Davies M (2004) Identification of a residue in the gamma-aminobutyric acid type A receptor alpha subunit that differentially affects diazepam-sensitive and -insensitive benzodiazepine site binding. J Neurochem 88: 1431–1438.

50. Newell JG, McDevitt RA, Czajkowski C (2004) Mutation of glutamate 155 of the GABA$_A$ receptor beta2 subunit produces a spontaneously open channel: a trigger for channel activation. J Neurosci 24: 11226–11235.

51. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410.

52. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52: 696–704.

53. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

54. Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21: 2104–2105.

55. Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8: 275–282.

56. Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39: 306–314.

57. Huelsenbeck JP, Larget B, Alfaro ME (2004) Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. Mol Biol Evol 21: 1123–1133.

58. Yang Z (2005) Bayesian inference in molecular phylogenetics. In: Gascuel O, ed. Mathematics of evolution and phylogeny. Oxford: Oxford University Press. pp 63–90.

59. Bielawski JP, Yang Z (2003) Maximum likelihood methods for detecting adaptive evolution after gene duplication. J Struct Funct Genomics 3: 201–212.

60. Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 13: 555–556.

61. Nielsen R, Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148: 929–936.

62. Venclovas C, Margelevicius M (2005) Comparative modeling in CASP6 using consensus approach to template selection, sequence-structure alignment, and structure assessment. Proteins 61 Suppl 7: 99–105.